

When cheaper is better: fee determination in the market for equity mutual funds

Gil-Bazo, Javier; Ruiz-Verdú, Pablo

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:

Gil-Bazo, J., & Ruiz-Verdú, P. (2008). When cheaper is better: fee determination in the market for equity mutual funds. *Journal of Economic Behavior & Organization*, 67(3-4), 871-885. <https://doi.org/10.1016/j.jebo.2007.04.003>

Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu>. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

gesis
Leibniz-Institut
für Sozialwissenschaften

Terms of use:

This document is made available under the "PEER Licence Agreement". For more information regarding the PEER-project see: <http://www.peerproject.eu>. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der

Leibniz-Gemeinschaft

Accepted Manuscript

Title: When Cheaper is Better: Fee Determination in the Market for Equity Mutual Funds

Authors: Javier Gil-Bazo, Pablo Ruiz-Verdú

PII: S0167-2681(08)00030-9
DOI: doi:10.1016/j.jebo.2007.04.003
Reference: JEBO 2170



To appear in: *Journal of Economic Behavior & Organization*

Received date: 11-4-2005
Accepted date: 27-4-2007

Please cite this article as: Gil-Bazo, J., Ruiz-Verdú, P., When Cheaper is Better: Fee Determination in the Market for Equity Mutual Funds, *Journal of Economic Behavior and Organization* (2007), doi:10.1016/j.jebo.2007.04.003

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

When Cheaper is Better: Fee Determination in the Market for Equity Mutual Funds

Javier Gil-Bazo

*Universidad Carlos III de Madrid, Department of Business Administration,
Calle Madrid, 126. 28903, Getafe (Madrid), Spain*

Pablo Ruiz-Verdú *

*Universidad Carlos III de Madrid, Department of Business Administration,
Calle Madrid, 126. 28903, Getafe (Madrid), Spain*

Abstract

In this paper, we develop a model of the market for equity mutual funds that captures three key characteristics of this market. First, there is competition among funds. Second, fund managers' ability is not observed by investors before making their investment decisions. Third, some investors do not make optimal use of all available information. The main results of the paper are that 1) price competition is compatible with positive mark-ups in equilibrium, and 2) worse-performing funds set fees that are greater or equal to those set by better-performing funds. These predictions are supported by available empirical evidence.

Keywords: Mutual fund fees, mutual fund performance, product quality, asymmetric information, bounded rationality.

JEL classification: L13, L15, G23, D80.

* Corresponding author. Tel.: +34 916245801; fax: +34 916249607.

Email addresses: javier.gil.bazo@uc3m.es (Javier Gil-Bazo), pablo.ruiz@uc3m.es (Pablo Ruiz-Verdú).

1 Introduction

In 2005, total mutual fund assets in the U.S. were worth 8.9 trillion dollars and constituted an estimated 20 percent of the total financial wealth of U.S. households (Investment Company Institute 2006). The increasing reliance of American investors on mutual funds has raised concerns among industry commentators and regulators alike about the extent of competition in the industry and, in particular, about the level of fees that mutual funds charge investors for their services.¹ These concerns have prompted regulators to commission reports (Securities and Exchange Commission 2000, General Accounting Office 2000) to analyze the evolution and determinants of mutual fund fees. More recently, a series of scandals² have triggered regulatory initiatives that, among other things, strengthen mutual fund fee disclosure requirements so as to promote price competition in the industry.³ Underlying this debate is a long overdue question: How are fees determined in the mutual fund market? To address this question, in this paper, we develop a model of the market for equity mutual funds.

One of the main concerns fueling the debate over mutual fund fees, and a key motivation for the regulatory changes requiring improved fee disclosure, has been the degree to which investors are aware of the fees associated with fund investments and their impact on the return of those investments. A survey by the SEC and the Office of the Comptroller of the Currency (Alexander et al. 1997) reports that fewer than one in six fund investors understand that higher fund expenses, which are deducted from the fund's assets and mostly consist of management fees, lead to lower returns. The survey also documents that investors are not familiar with the level of fees they are paying for mutual fund services (not even 20% of the respondents was able to give an estimate of the expenses paid for their largest mutual fund) and reveals other gaps in financial literacy. Similar results have been obtained in studies by Capon et al. (1996) and Wilcox (2003). More recently, a study by Barber et al. (2005) has provided further evidence of investors' difficulties in understanding the effects of mutual fund fees. In light of these findings, it is our view that a satisfactory model of the mutual fund market should account for the presence

¹ For example, in a recent article about the topic in *The Economist*, one could read, "Retail investors [...] have seen precious little competition on prices. Even for wretched performance, reductions in fees have been all too rare" (*The Economist* 2004).

² See, for instance, Dwyer et al. (2003) for an account of these scandals.

³ The Securities and Exchange Commission has issued a rule establishing stronger disclosure requirements for mutual funds (Securities and Exchange Commission 2004). The "Mutual Funds Integrity and Fee Transparency Act of 2003", which imposes strict governance and disclosure requirements, passed the U.S. House overwhelmingly in November, 2003.

of a significant fraction of investors who make less than optimal use of the information available when making their investment decisions.

The model developed in this paper has three main ingredients, which, we believe, characterize the market for equity mutual funds. In the U.S., investors can choose from a large pool of mutual funds, even for relatively narrowly defined investment categories. Therefore, the first ingredient of our model is competition among (possibly many) funds for investors' money. This ingredient is key because we would like to know whether existing concerns about the level of fees charged by mutual funds can still be warranted in the presence of mutual fund competition. Since our main focus is the market for equity mutual funds, the second ingredient is quality uncertainty: different fund managers have different abilities to generate returns, but those abilities are not known *ex ante* by investors. As we argue below, even though in actual markets, funds' past returns could be used as a signal of performance, this signal is, at best, noisy. Therefore, the assumption that quality is not observable appears as a reasonable first approximation. The third ingredient is the presence of a fraction of investors who do not make optimal use of all available information when making their investment decisions. We label these investors *unsophisticated*.

The model enables us to address several questions. 1) Should we expect price competition among funds to bring equilibrium profits down to zero? 2) Will high-quality funds drive low-quality funds out of the market? If low-quality funds manage to survive in equilibrium, 3) will high-quality funds charge higher fees, so that, in equilibrium, all investors earn the same expected net returns? If, on the contrary, net returns differ across funds, 4) which funds will be more likely to overcharge investors: high- or low-quality funds? These questions lie at the heart of the debate over the extent of competition in the mutual fund industry and are of key importance in evaluating proposed regulatory changes. The model shows that, in equilibrium, 1) funds earn positive profits, 2) high- and low-quality funds coexist, 3) high-quality funds never charge higher fees and may charge lower fees, and that, as a consequence, 4) low-quality funds greatly overcharge investors.

The intuition behind these (perhaps surprising) results is as follows. In the mutual fund market, the revenues earned by a mutual fund are the product of its fee and the fund asset value. Therefore, in a setting in which quality is unobservable, high-quality funds, namely those that deliver a higher value, may be able to differentiate themselves by setting low fees. If those fees are low enough, low-quality funds would not break even by imitating them. Unable to compete for the sophisticated segment of the market, low-quality funds would focus instead on extracting rents from unsophisticated investors. It should be noted that, although the presence of unsophisticated investors is necessary to obtain this sort of price differentiation in equilibrium, it is not sufficient. The

model's results follow from the interaction between asymmetric information and the existence of those investors.

The assumption that fund quality is completely unobservable can seem unrealistic if, as indicated by the empirical evidence, past mutual fund returns have some value as predictors of future performance. The complexity involved in evaluating and forecasting mutual fund performance, however, suggests that past returns are only a noisy signal of future performance. The assumption of asymmetric information is, therefore, still reasonable given that managers possess skills, technology and access to return-relevant data that investors lack. In section 4, we investigate the effects of allowing investors to observe a signal of fund quality prior to making their purchase decision. Our analysis shows that, as long as the signal is only imperfectly correlated with fund performance, separating equilibria exist and are like the ones described above.

The available empirical evidence is consistent with our model: funds not only fail to adjust fees so as to offset differences in before-fee returns, but on the contrary, funds of lower quality seem to charge higher fees. Elton et al. (1993), for instance, document significant differences in after-fee returns across mutual funds and show that funds that charge higher fees deliver significantly lower before-fee returns. Additional evidence provided by Gruber (1996), Carhart (1997), Harless and Peterson (1998), and Chevalier and Ellison (1999) has consistently confirmed the power of high fees to predict underperformance. Surprisingly, this anomalous pattern in the data has not been explained by any previous formal model of the mutual fund industry.

Although there exists a relatively large theoretical literature, initiated by Bhattacharya and Pfleiderer (1985), that characterizes the optimal compensation contract in a delegated portfolio management problem,⁴ few studies have analyzed fund fees as the outcome of the strategic interaction of competing mutual funds. Recently, Hortacsu and Syverson (2004) have developed a search model of the market for S&P 500 index funds. In contrast to our paper, however, they analyze a sector in which financial performance differences across funds are relatively small and thus focus on non-portfolio fund differentiation and search frictions as potential sources of fee dispersion. In another recent paper, Berk and Green (2004) have proposed a model of the mutual fund market with no informational asymmetries or search frictions. The goal of their model is to explain why money may rationally follow past good performance even when past performance is not a strong predictor of future performance. The paper is, however, silent about the determinants of the observed distribution of fees. Das and Sundaram (2002) and Metrick and Zeckhauser (1999) have analyzed fee setting in a duopoly context with asymmetric information about fund quality. While Das and Sundaram compare the performance of

⁴ See, for example, Palomino and Prat (2003) for a recent contribution.

two types of incentive schemes for fund managers in such a context, the goal of Metrick and Zeckhauser is to explain why high- and low-quality producers may charge the same price in certain markets (including the mutual fund market), and is, thus, closely related to ours. We discuss their results in section 2. In another related paper, Nanda et al. (2000) have developed a model where mutual fund managers of observable quality bear the cost of stochastic investment redemptions and, therefore, wish to attract investors who are less likely to experience liquidity needs. In equilibrium, more skilled fund managers impose exit fees and cater to investors with lower liquidity needs, while less skilled managers become liquidity providers. Finally, Christoffersen and Musto (2002) have explored empirically the effect of investors' performance sensitivity on fund fees.

Our paper is also related to the more general literature on the role of prices as signals of quality. In this literature, high prices generally signal high quality (e.g., Bagwell and Riordan 1991), although in some contexts involving repeated purchases, it has been shown that low introductory prices can be used as signals of quality (Schmalensee 1978). This set of work, however, has mostly focused on the case of a single seller of unknown quality. Only recently, Fluet and Garella (2002) and Hertzendorf and Overgaard (2001) have studied price and advertising signaling for the case of a duopoly. In their models, high-quality firms charge higher prices in any separating equilibrium. Of related interest are papers by Gabaix and Laibson (2003, 2006), in which the authors show that product—or pricing—complexity can allow firms to obtain positive mark-ups in equilibrium and that increases in competition may exacerbate the problem.

The rest of the paper is organized as follows: section 2 presents the benchmark model, section 3 extends the model to include unsophisticated investors, section 4 assesses the consequences of the existence of a noisy signal of fund quality, section 5 discusses the available evidence on the relationship between performance and fees, and, finally, section 6 concludes.

2 A Model of Fee Determination in the Market for Mutual Funds

Consider a simple setting in which there is a continuum of investors of mass one who have one dollar to invest and N mutual fund managers. These managers can be of one of two types depending on their ability: good (g) and bad (b). G-managers earn gross expected return R_g , and b-managers R_b , where $R_g > R_b$, and $R_g > 1$. Returns are assumed to be independent of fees, so we can abstract from moral hazard problems. The *ex ante* distribution of types is given by the probability p that a manager is good. Once the types are realized, fund managers observe their quality but not the quality of their

rivals and decide what fraction e of the fund's final asset value to charge to investors. The assumption that fees are determined as a fraction of asset value is made to reflect actual practice in the mutual fund industry. In section 5, we discuss in greater detail the different fees paid by mutual fund investors.

Investors do not observe quality, so they decide where to invest on the basis of the prior distribution and the fees charged by the different funds. As argued in the introduction, even if fund quality could be inferred from different sources of information, such as past returns, evaluating and forecasting mutual fund performance is a very complex and costly task for retail investors. Fund managers, on the other hand, possess skills, technology and access to data about their portfolios and organization that investors lack. We therefore believe that fund managers can be expected to be better informed about their own ability to generate returns than investors. We discuss this issue in greater detail in section 4, where we allow investors to be imperfectly informed about fund quality.

We will assume that a fraction γ of all investors are unsophisticated, in the sense that they do not make optimal use of all available information. Since we will first analyze the case with no unsophisticated investors, we postpone the discussion of the precise form in which these investors behave until they are introduced in section 3.

The costs of managing the fund are cw , where w is the amount of money managed by the fund manager. It is assumed that costs are low enough to make it profitable for g-funds to operate if their type is known. We assume that all market participants are risk-neutral,⁵ and that the only alternative investment is a risk-free asset paying a zero interest rate. Therefore, the maximum fee a fund of type k can charge if its type is known is such that the net return for investors is equal to one: $R_k(1 - e) = 1$, that is $e = \frac{R_k - 1}{R_k}$. Therefore, the assumption that g-funds may profitably operate if their type is known can be expressed as follows:

⁵ The risk-neutrality assumption enables us to focus on a single mutual fund characteristic, expected return, so rational investors will always choose the fund with the highest (after-fee) expected return. We could alternatively drop risk neutrality and replace expected return with a measure of mutual fund risk-adjusted performance. A commonly used proxy for risk adjusted-performance is the fund's alpha, defined as the fund's expected return in excess of that implied by the Capital Asset Pricing Model (CAPM) or an Arbitrage Pricing Theory (APT) -type model. Under both the CAPM and the APT assumptions, investors would prefer funds with higher (after-fee) alpha.

Assumption 1

$$\frac{R_g - 1}{R_g} > \frac{c}{R_g},$$

where the right-hand side of the inequality is the break-even fee for g-funds.

We also assume that, given c , b-funds may find it profitable to operate for some fee less than one hundred percent:

Assumption 2 $R_b > c$.

We denote by e_k the fee charged by a manager of type k as a proportion of the value of the fund at the end of the period and assume that there are no other fees. Therefore, the amount paid by an investor who invests w dollars in a fund of type k is $w e_k R_k$, and payoffs are $w(e_k R_k - c)$ for the manager and $w(1 - e_k)R_k$ for the investor.

Finally, the timing of decisions is as follows. First, managers simultaneously set fees. Then investors decide where to invest. We make the assumption that, if several funds have the same net expected returns, investors allocate their wealth among them with equal probability.

2.1 Benchmark Case: Complete Information

Before solving the model, it is instructive to investigate the relationship between fund quality and fees when there are no unsophisticated investors and quality is observed both by competing funds and by investors. It is straightforward to show that, in this case, b-funds will be driven out of the market whenever there are g-funds:

Proposition 1 *With complete information, there do not exist equilibria in which both fund types operate simultaneously.*

Proof. First, note that for both types of funds to have a positive market share,

$$(1 - e_b)R_b = (1 - e_g)R_g, \tag{1}$$

which implies that $e_b < e_g$.

With complete information the number of funds of each type is given and commonly known. We study first the case in which there is only one g-fund active and then the case in which there are at least two g-funds. So, suppose

first that there is only one g-fund. For b-funds to operate $e_b \geq \frac{c}{R_b} > \frac{c}{R_g}$, which implies that $e_g > \frac{c}{R_g}$. Therefore, if $e_b \geq \frac{c}{R_b}$, it would be profitable for the g-fund to lower e_g slightly and attract the whole market. It follows that, in equilibrium, b-funds cannot operate. The g-fund will set the minimum fee that guarantees that b-funds do not want to enter, which is given by:

$$(1 - e_g)R_g = \left(1 - \frac{c}{R_b}\right) R_b. \quad (2)$$

If there are several g-funds, the same argument applies for any $e_g > \frac{c}{R_g}$. The only possible equilibrium fee is $e_g = \frac{c}{R_g} < \frac{c}{R_b}$, so b-funds remain inactive. ■

Therefore, good and bad funds cannot coexist in equilibrium with complete information: whenever it is profitable for b-funds to operate, it is also profitable for g-funds to lower fees.⁶ As a result, b-funds are driven out of the market.

2.2 Asymmetric Information

We now investigate what happens when there is asymmetric information regarding funds' ability to generate returns. To understand the contribution of the model's different assumptions, we maintain the assumption that there are no unsophisticated investors. We use the Perfect Bayesian Equilibrium as our equilibrium concept and focus only on pure-strategy symmetric equilibria (*i.e.*, equilibria in which all funds of the same type play the same pure strategy). To limit equilibrium multiplicity, we require investors' out-of-equilibrium beliefs to satisfy the property that they do not assign positive probability to managers setting fees that are certain to yield them a negative profit. That is, investors cannot assign a positive probability to a fund of type k choosing a fee less than $\frac{c}{R_k}$. Therefore, throughout the paper, by *equilibrium* we will refer to a pure-strategy Perfect Bayesian Equilibrium satisfying this restriction on investors' beliefs.

First, note that in a separating equilibrium in which both types are ever active simultaneously, it has to be the case that net returns for investors are equal across types since, otherwise, investors would not invest with the two types when both are available. Equality of net returns, in turn, implies that

⁶ In models of vertical differentiation (*e.g.*, Shaked and Sutton 1982), equilibria in which low- and high-quality producers coexist and in which the former charge lower prices are possible if consumers display differences in their willingness to pay for quality. In the mutual fund industry, however, where the good provided by sellers is end-of-period *dollars*, one would expect that all consumers have the same willingness to pay for quality: nobody would pay more cents than anybody else for a dollar.

the expected market shares of g- and b-funds also have to be equal,⁷ since investors are indifferent between both types and, therefore, allocate the money across funds with equal probability. However, this implies that, if $e_g > e_b$, it would be optimal for b-managers to imitate g-managers. On the other hand, if $e_g < e_b$, no rational investor who observes both fees would invest with a b-manager. Therefore,

Proposition 2 *If all investors react optimally to differences in expected pay-offs, there are no separating equilibria in which both fund-types operate simultaneously.*

According to Proposition 2, we should observe no fee dispersion in equilibrium: for any realization of the number of b- and g-funds, all the funds with a positive market share must charge the same fees. Equilibria at which b- and g-funds are active simultaneously and the latter charge higher fees are not possible.

It is straightforward to show that, with asymmetric information, equilibria such as the one obtained with complete information, in which only g-funds are active and make zero profits, are not possible. The reason is that any g-fund faces a positive probability of competing only against b-funds. Therefore, there exists a strategy that guarantees positive expected profits to g-funds: setting a fee greater than the break-even fee for g-funds ($\frac{c}{R_g}$), but lower than the break-even fee for b-funds ($\frac{c}{R_b}$). Such a fee guarantees positive profits if the fund is ever able to attract any money and, as long as it is not too high, ensures that the fund would attract investors' money at least when there are no competing g-funds. The fact that g-funds have a strategy that guarantees positive expected profits immediately rules out Bertrand-like equilibria, such as the one that results from complete information. It is possible to show, though, that, for a broad range of parameter values, there are equilibria in which both types of funds are active, set the same fee, and obtain positive profits. With complete information, equilibria with positive profits are not possible. If there is asymmetric information, however, undercutting other funds' fees can be interpreted by investors as a signal of low quality, so that, if investors are sufficiently pessimistic in their assessment of deviating funds, equilibria with positive profits can arise.⁸

We would like to note here that the existence of an equilibrium at which funds of different qualities set the same fee has already been proposed by Metrick and Zeckhauser, although in a very different context. They study a vertically-differentiated duopoly characterized by sequential price setting (with good funds setting fees, front-end loads, before bad funds) and by investor hetero-

⁷ The expectation is taken over the possible realizations of the number of b- and g-funds.

⁸ We omit the proof of the existence of this type of equilibrium, since it can be obtained as a special case of the pooling equilibria analyzed in the next section.

geneity along two dimensions: on the one hand, different investors value the “good” provided by mutual funds differently, and on the other hand, some investors can observe quality while others cannot. In this context, an equilibrium in which both funds set the same price can arise when the qualities are similar enough. The reason is that competition for the investors who can observe quality is strong in this case. As a result, the good fund may find it optimal to set a fee low enough to force bad funds out of the informed segment of the market. It is important to note that, in their model, good funds attract more money than bad funds in a pooling equilibrium, since the latter do not get any money from informed investors. It is also worth noting that, in their model, there are also separating equilibria in which both funds are active and good funds charge higher fees. In our model, these equilibria are not possible.

3 Unsophisticated Investors

In the introduction, we argued that a satisfactory model of the mutual fund industry should account for the empirical evidence that a significant fraction of investors do not make optimal use of all available information when making their investment decisions. For example, in a survey conducted by Capon et al. (1996), more than 70% of respondents did not know whether their mutual funds invested mainly in domestic or international stocks, or whether they invested mainly in equity or fixed income. Especially notable is investors’ ignorance of the fees charged by their mutual funds. In particular, a survey commissioned by the SEC and the Office of the Comptroller of the Currency (Alexander et al. 1997) shows that 81.2 % of respondents could not give an estimate of the expenses of the largest fund they owned and, of those investors, only 43% claimed to have known their largest fund’s expenses at the time they first invested in the fund. Investors’ ignorance of management fees is consistent with the importance that they assign to these fees when determining their investment choices: Capon et al. report that the mean importance given by individual investors to management fees as a selection criterion is 2.28 in a scale from 1 to 6, and Wilcox finds similar results when inferring investors’ selection criteria from actual choices. If investors weight any fees at all, they consider load fees and not management fees (Barber et al.).

Investors’ insensitivity to management fees may reflect the lack of salience of these fees (Barber et al.): management fees are paid over time, are calculated as a fraction of the value of the investment, and are, therefore, hard to translate into (discounted) dollar terms. Further, since expenses are typically in the range of 1-2%, some investors may underestimate the long run effects of apparently small differences in expenses or code as equal expenses with different long run effects. The problem of processability of information on mutual fund expense ratios is stressed by Cox and Payne (2005), who apply Slovic’s (1972)

concreteness principle to argue that investors tend to accept the information on fees in the format it is given rather than expend cognitive effort to transform it, thus providing a justification to the SEC's new rule requiring funds to disclose expenses in dollar terms. However, Cox and Payne also argue that these disclosure requirements may not be sufficient and that management fees will be underweighted in decisions unless they are expressed relative to those of comparable funds.

The evidence, thus, shows that the behavior of a large fraction of investors is markedly different from that of the sophisticated investors we considered in the previous section, who not only knew the fees charged by all funds, but rightly understood in equilibrium the relation between fees and expected performance. In this section, in addition to sophisticated investors, we include a fraction of investors whose investment choices are largely unresponsive to fees. More specifically, we assume that the probability of an unsophisticated investor investing in a given mutual fund is the same for all funds and independent of the fee charged by the fund as long as that fee is below a threshold value e_U . This threshold reflects the fact that, even though investors may not be fully aware of the exact value of fees or of their implications for net returns when they lie within a reasonable range, they would detect fees that are conspicuously high. We take a conservative approach and assume that this threshold is relatively low and given by:

$$\bar{R}(1 - e_U) = 1,$$

where $\bar{R} = pR_g + (1 - p)R_b$ is the expected gross return given the ex ante distribution of fund types. That is, unsophisticated investors invest only if the expected net return, given the distribution of fund types, is at least as high as that of the alternative asset. Arguably, this requirement demands too much sophistication from unsophisticated investors. As we discuss below, however, our results do not depend on the specific choice of e_U and are consistent with both higher and lower values of this parameter. Finally, we assume that unsophisticated investors do not infer any information about a fund's quality from the fee it charges. Again, we could somewhat relax this assumption. However, as we argue below, we do not think that doing this would increase the realism of our assumptions and would unnecessarily complicate the results.

To derive the market equilibrium when there is a fraction γ of unsophisticated investors, note first that a reasoning similar to Proposition 2 still applies in this case: in equilibrium, b-managers and g-managers cannot both serve the sophisticated market segment and charge different fees. If $e_g > e_b$, b-managers would mimic g-managers' pricing strategy. If $e_g < e_b$, sophisticated investors would not invest in b-funds. Therefore, if the presence of unsophisticated investors allows for the existence of separating equilibria in which both fund types are active simultaneously, sophisticated investors must prefer one type

of fund over the other. The possibility that b-funds offer a higher net return in equilibrium can be ruled out, since it is straightforward to show that there cannot exist separating equilibria in which g-funds serve only unsophisticated investors and b-funds serve sophisticated investors. Therefore, we investigate next whether there can exist separating equilibria in which both fund types are active simultaneously and in which b-funds serve only unsophisticated investors as long as there are competing g-funds. Obviously, for these equilibria to exist, b-funds must be able at least to break even by charging e_U since, otherwise, they would never be active.

Assumption 3 $e_U > \frac{c}{R_b}$.

The following conditions must hold at this type of equilibrium:

$$w_g^U(R_g e_g - c) \geq w_b^U(R_g e_b - c) \quad (\text{NIg})$$

$$w_b^U(R_b e_b - c) \geq w_g^U(R_b e_g - c) \quad (\text{NIb})$$

$$w_b^U(R_b e_b - c) \geq 0, \quad (\text{Pb})$$

where w_k^U is the wealth that a fund setting e_k expects to obtain conditional on all other funds playing the equilibrium strategies. Note that $e_g < e_b$ implies that if b-funds serve any sophisticated investors at all, they will do so only when there are no g-funds in the market, whereas g-funds serve sophisticated investors with probability one. Since all funds get the same share of unsophisticated investors' money, it follows that $w_g^U > w_b^U$.

The first two conditions above are no-imitation constraints for g- and b-funds, respectively, and the last condition is a participation constraint for b-funds. A participation constraint for g-funds is not necessary because it is implied by (NIb) and (Pb). Note that since $w_g^U > w_b^U$, condition (NIb) requires that $e_g < e_b$; in this type of equilibrium, g-funds must set lower fees.

Fees also have to be low enough to convince both sophisticated and unsophisticated investors to participate:

$$e_g \leq \frac{R_g - 1}{R_g} \quad (3)$$

$$e_b \leq e_U \quad (4)$$

Finally, it cannot be profitable for b- or g-funds to deviate and set an out-of-equilibrium fee. To evaluate these deviations, we need to make assumptions about sophisticated investors' out-of-equilibrium beliefs. To prove the results below, we assume that sophisticated investors interpret any deviation from equilibrium as coming from a b-fund unless it yields negative profits for such a fund. These extreme beliefs are chosen to simplify the proofs and are not necessary. For our results to hold, all that is required is that sophisticated

investors assign a sufficiently high probability to a deviator being of low quality. The next proposition shows that there are parameter values such that all the above conditions hold simultaneously and there are no profitable out-of-equilibrium deviations (proofs can be found in the appendix).

Proposition 3 *There exist separating equilibria with unsophisticated investors at which:*

- (1) *b-funds serve unsophisticated investors only and charge $e_b^* = e_U$,*
- (2) *g-funds charge $e_g^* > \frac{c}{R_b}$ and serve both sophisticated and unsophisticated investors,*
- (3) *$e_b^* > e_g^*$.*

Figures A.1 and A.2 show that separating equilibria of this sort can exist for reasonable parameter values. For each possible value of γ (plotted along the x-axis), the figures graph the minimum and maximum values of c (plotted along the y-axis) for which these equilibria can exist.

The equilibria described in Proposition 3 are the only possible separating equilibria, but there may also exist pooling equilibria for certain parameter values. At these pooling equilibria, g- and b-funds set the same fee, and both types make positive profits.

Proposition 4 *For some parameter values, there exist pooling equilibria in which both types set $e_p > \frac{c}{R_b}$.*

Given the complexity of the conditions that define both the pooling and the separating equilibria, we have not attempted to ascertain which type of equilibrium holds for a greater range of parameter values or investors' beliefs. As we did above for the case of separating equilibria, we have numerically found the range of values of c and γ consistent with the existence of a pooling equilibria for the case in which $p = \frac{1}{2}$. Figures A.3 and A.4 show that pooling equilibria can also exist for reasonable parameter values.

Inspection of propositions 3 and 4 seems to indicate that quality is not rewarded in the market for mutual funds since g-funds charge fees that are lower or equal to those charged by b-funds. Quality is, however, rewarded by higher profits: at any equilibrium, g-funds obtain higher profits than b-funds. At a separating equilibrium, the no-imitation constraint (NIg), ensures that $w_g^U(R_g e_g - c) \geq w_b^U(R_g e_b - c)$, which, since $R_g > R_b$, implies that $w_g^U(R_g e_g - c) > w_b^U(R_b e_b - c)$. At a pooling equilibrium, all funds charge the same fee, e_p and get the same market share. The fact that $R_g > R_b$ immediately implies that $\frac{1}{N}(R_g e_p - c) > \frac{1}{N}(R_b e_p - c)$.

The model in this section departs from the benchmark complete information model in two dimensions, and it is instructive to see how each of these di-

mensions contributes to the existence of separating equilibria such as the ones described in Proposition 3. First, the existence of unsophisticated investors allows b-funds to survive while setting fees that differ from those of g-funds. As we saw in the previous section, this would not be possible if all investors held correct beliefs in equilibrium and could move their money freely. In this respect, our model resembles models of price dispersion based on the presence of search costs (like Salop and Stiglitz 1977), where impediments to search allow firms charging higher prices to obtain positive market shares. Second, the presence of asymmetric information limits the competitive pressure on g-funds. If sophisticated investors could observe fund quality, competition among g-funds would drive e_g down to $\frac{c}{R_g}$, but such a situation could not be an equilibrium with unsophisticated investors, because g-funds can set a higher fee, sell to unsophisticated investors and make a positive profit. It should thus be emphasized that the existence of unsophisticated investors *alone* cannot generate separating equilibria with both fund types active.

In light of Proposition 3, it is worth reevaluating whether our characterization of unsophisticated investors implies that these investors behave in a way that would be considered unreasonable even for investors such as those described by the survey evidence discussed at the beginning of this section. A possible concern is that

$$e_U > \frac{R_b - 1}{R_b},$$

where $\frac{R_b - 1}{R_b}$ is the fee that guarantees the reservation return when investing with a b-fund. Therefore, at a separating equilibrium, some unsophisticated investors (those paired with b-funds) would do better by investing in the reservation asset. We do not think that this requires too much unsophistication from investors for two reasons. First, unsophisticated investors earn an average return at least as high as that of the alternative asset, since a fraction p of unsophisticated investors are paired with g-funds. The average return for unsophisticated investors is

$$\begin{aligned} p(1 - e_g^*)R_g + (1 - p)(1 - e_U)R_b &> p(1 - e_U)R_g + (1 - p)(1 - e_U)R_b = \\ &= (1 - e_U)(pR_g + (1 - p)R_b) = (1 - e_U)\bar{R} = 1. \end{aligned} \quad (5)$$

The second reason is that most studies agree that the average actively managed mutual fund has historically delivered *below* market returns, at least after expenses and transaction costs are deducted (see, e.g., Wermers 2000). In this sense, our model would seem to require too much (rather than too little) sophistication from investors.

Another possible concern relates to the assumption that unsophisticated investors do not update their beliefs about funds' quality based on the fees charged by those funds. Given that, in a separating equilibrium, good and

bad funds set different fees, one may wonder whether separating equilibria such as the ones described above could survive. There are again two reasons why we are not especially worried by this concern. The first reason is theoretical. Although we have assumed that e_U is the maximum fee that an investor would like to pay for a fund of average quality, the proof of Proposition 3 shows that a separating equilibrium can exist as long as $e_U > 1 - \frac{1}{R_b}$, that is, as long as the maximum fee that unsophisticated investors are willing to pay is greater than the maximum fee a sophisticated investor would be willing to pay for a low-quality fund. Therefore, as long as unsophisticated investors do not fully update their beliefs (*i.e.*, if they do not believe that a fund setting e_U is bad with probability one), separating equilibria can exist. The second reason is empirical. As we discuss in section 5, a negative relation between expenses and performance has been extensively documented. In the survey reported by Alexander et al., however, about 20 percent of respondents believed that mutual funds with higher expenses produced better results, 64.4 percent believed that funds with higher expenses produced average results, and only 15.7 percent of the survey respondents believed that higher expenses led to lower-than-average returns. Since the negative relationship between fees and performance had been already identified by several studies before the survey was conducted, the survey shows that a significant fraction of investors had not optimally updated their beliefs. Further, investors' money, and not just investors' opinion, seems to show that a sizeable number of them are not making the right investment choices. As Martin Gruber put it, we see that money remains in funds that can be predicted to do poorly and that, in fact, do perform poorly (p. 807).

4 Noisy Performance Signals

Throughout the article, we have assumed that while fund managers knew their expected performance, the only information available to investors prior to their investment decision was contained in the fees charged by mutual fund managers. If funds differ in their ability to generate returns, however, we should expect past returns to convey some information about fund quality. Therefore, we would like to know if our results are robust to the existence of a noisy signal of fund quality. Further, if past returns were accurate predictors of future performance, the very assumption of asymmetric information could be put into question, so it is important to know to what extent future performance can be predicted on the basis of information about past returns.

The empirical evidence shows that significant underperformance is a reasonably good predictor of future underperformance. In other words, there seems to exist performance persistence for worst-performing funds. The existence of persistence at the top, however, remains a hotly debated issue. Grinblatt and

Titman (1992) and Hendricks et al. (1993), among others, find that previous years' winners and losers tend to repeat (this is known as the "hot hand" phenomenon). Carhart, however, questions these findings, explaining the previously documented persistence of top performers in terms of survivorship bias and the momentum anomaly in the stock market. More recent evidence (see, for instance, Bollen and Busse, 2005), however, shows that there is some performance persistence at the top, although this persistence is short lived (not longer than a quarter).

Two conclusions can be drawn from this evidence. First, information on past returns has some predictive value, although this predictive value is limited and, second, as evidenced by the lack of consensus in the literature, the problem of predicting future performance on the basis of past returns is very complex and requires advanced financial and quantitative skills. Taken together, these conclusions imply that for individual investors, past returns will be, at best, a noisy signal of performance. The limited information conveyed by past returns also implies that managers may be better informed than investors. In other words, the observed performance persistence is compatible with the existence of asymmetric information. Moreover, although the fact that fund managers may be better informed than investors does not justify assuming that they will, in fact, be better informed, we believe that one should expect fund managers to have a better assessment than investors about their expected performance. The reason is twofold. First, fund managers are likely to possess more advanced financial skills and better research tools than individual investors. Second, and most importantly, fund managers have access to return-relevant information not available to investors. For instance, managers continuously observe the composition of the portfolio they manage, while investors can observe only the fund's portfolio composition at some pre-specified moments throughout the year. Fund managers also observe each trade they make and the transaction costs associated with each of those trades. Further, managers are likely to have a better idea of their skills, the skills of their staff, and the quality of their research tools. In particular, fund managers can assess the accuracy of their own forecasts and, thus, assess whether fund results are due to their investment strategy or to luck. Finally, fund managers are better informed than investors about the likelihood of future changes in the fund's staff (in particular, the probability of their own departure from the management team) or investment strategy.

In our view, the limited value that past returns have for investors as predictors of future performance justifies making the simplifying approximation that there is, in fact, no information available to them other than fund fees. To check the robustness of our results, however, we next extend the model to allow investors to observe a noisy signal of performance before making their investment decision. In the interest of brevity, we do not analyze pooling equilibria and focus strictly on whether the results obtained in the previous section con-

cerning the relationship between fees and fund quality at separating equilibria still hold when there is a noisy signal of performance.

Suppose that after funds set fees, investors observe a noisy signal σ of fund quality, which can take two values, l or h (low or high). Conditional on the fund's type, σ is i.i.d. across funds with a probability distribution given by $Pr(l|b) = Pr(h|g) = \nu > 1/2$. Therefore,

$$Pr(g|h) = \frac{p\nu}{p\nu + (1-p)(1-\nu)} \equiv h > p, \quad (6)$$

$$Pr(g|l) = \frac{p(1-\nu)}{p(1-\nu) + (1-p)\nu} \equiv l < p. \quad (7)$$

That is, upon observing h the probability of a fund being g is $h > p$, and, upon observing l the probability of a fund being g is $l < p$. The higher ν , the higher h and the lower l and, thus, the higher the precision of the signal.

While sophisticated investors will optimally condition their investment decision on performance signals and fees, the response of unsophisticated investors to the presence of a performance signal is less clear a priori. On the one hand, unsophisticated investors may not respond, or respond too little, to the signal if they do not observe it or underestimate the amount of information carried by it. On the other hand, unsophisticated investors may overreact to the signal if they overestimate its precision. To capture the whole possible range of behavior by unsophisticated investors we assume that the fraction of the wealth of unsophisticated investors that a fund with a value j ($j = l, h$) for the signal expects to obtain is given by a function f_j of the other funds' signals. By varying this function we can account for any degree of responsiveness to the signal by unsophisticated investors. For example, assuming that $N = 2$, if unsophisticated investors do not respond to the signal, then $f_h(\sigma_h) = f_h(\sigma_l) = f_l(\sigma_h) = f_l(\sigma_l) = \frac{1}{2}$ (where the argument of the function in each case is the realization of the other fund's signal). If unsophisticated investors' investment choice is fully determined by the signal, then $f_h(\sigma_h) = f_l(\sigma_l) = \frac{1}{2}$, $f_h(\sigma_l) = 1$, and $f_l(\sigma_h) = 0$.

Independently of how unsophisticated investors respond to the signal, it is still the case that if there exist separating equilibria, then worse-performing funds charge higher fees since at a separating equilibrium, fees are perfect signals of fund quality, while σ is only a noisy signal. Therefore, sophisticated investors will optimally disregard the signal and infer quality from the fee alone. As in previous sections, it follows that if $e_g > e_b$, then it would be profitable for b-funds to imitate g-funds. Therefore, at any separating equilibrium, $e_g < e_b$.

To check whether separating equilibria such as the ones described in Proposition 3 exist in the presence of a signal of fund quality, suppose first that unsophisticated investors disregard the signal σ . In this case, it is straight-

forward to check that the conditions that guarantee existence of a separating equilibrium are exactly identical to the ones derived in the absence of σ .

To analyze the case in which unsophisticated investors condition their investment decision on the signal, let \bar{s}_k denote the money that a fund of type k expects to obtain from sophisticated investors at a separating equilibrium and \bar{u}_k the money that it expects to obtain from unsophisticated investors. With these definitions, we can write the no-imitation conditions at a separating equilibrium as follows:

$$\begin{aligned} (\bar{s}_g + \bar{u}_g)(R_g e_g - c) &\geq \bar{u}_g(R_g e_U - c), & (\text{NIg''}) \\ \bar{u}_b(R_b e_U - c) &\geq (\bar{s}_g + \bar{u}_b)(R_b e_g - c), & (\text{NIb''}) \end{aligned}$$

where the first condition ensures that g-funds do not want to set e_U , and the second one ensures that b-funds do not want to pass as g-funds. In both conditions, the money that a fund setting e_U expects to obtain from sophisticated investors at a separating equilibrium is zero, since $e_U > \frac{R_b - 1}{R_b}$. Note now that if $P(h|g) = \nu > 1 - \nu = Pr(h|b)$, then $\bar{u}_g > \bar{u}_b$, since a fund with $\sigma = h$ will obtain more money from unsophisticated investors in expectation than a fund with $\sigma = l$, and g-funds are more likely to have a high signal.

To characterize fully a separating equilibrium, we need to check the two no-imitation conditions above, participation constraints for b-funds and sophisticated investors, and constraints that ensure that funds do not want to deviate to a fee that is not played in equilibrium (that is, a fee other than e_g or e_U). These conditions are analogous to the ones derived in the absence of a signal, so we do not rewrite them here. Figures A.5, A.6, and A.7 display the range of parameter values for which separating equilibria exist for the case $N = 2$ and for different values of ν , f_l and f_h .⁹ Figure A.5 displays the range of values of c and γ for which separating equilibria exist when the signal is only slightly informative about the fund's type. In particular, ν is chosen to be 0.55. In each figure, we consider three different cases. First, unsophisticated investors ignore the signal; that is, $f_h(\sigma_h) = f_l(\sigma_l) = f_h(\sigma_l) = f_l(\sigma_h) = 0.5$. In this case, represented by the thick solid line, separating equilibria are identical to the ones in the absence of a signal (since unsophisticated investors are equally likely to invest with either fund). Second, unsophisticated investors are more likely to invest in a fund with high signal if the other fund has low signal. In particular, the thin solid line represents the parameter set compatible with equilibrium when $f_h(\sigma_l) = 0.8$ and $f_l(\sigma_h) = 0.2$. Finally (dashed lines), we consider the case in which, if the funds have different signals, unsophisticated

⁹ Note that in the figures, $R_b < 1$. This assumption greatly simplifies the computation of the equilibrium conditions and is empirically accurate: sophisticated investors would never invest with bad actively managed funds if they knew their quality. The authors would be glad to provide the program used to derive the figures upon request.

investors invest only in the fund with a high signal: $f_h(\sigma_l) = 1$ and $f_l(\sigma_h) = 0$. As one could have expected, the range of parameters for which a separating equilibrium exists shrinks as unsophisticated investors become more likely to invest with funds with a high value of the signal and as the precision of the signal increases, although the reduction is not severe.

Figures A.6 and A.7 show what happens as we increase the precision of the signal to $\nu = 0.7$ and $\nu = 0.9$, respectively.¹⁰ The figures show that separating equilibria still exist for reasonable parameter values in the limit, that is, as the signal becomes extremely precise, even if the investment decision of unsophisticated investors is completely determined by the signal. We can, therefore, conclude that the results derived in the previous section are largely robust to the presence of a noisy signal of performance. It should be noted that if $\nu = 0.50$, so that the signal is not truly informative, then separating equilibria would be identical to those derived in the absence of a signal, independently of the degree to which unsophisticated investors are affected by the signal.

It is worth noting that one could, alternatively, assume that σ is observed by funds and investors *before* funds set their fees. This alternative assumption would allow us to model not just how investors react to noisy performance information, but also how funds strategically respond to information about their own and other funds' performance. The problem with this assumption is that if the signals became public before funds set fees, a fund's decision would have to be made contingent not only on its quality and the value of its signal, but also on the complete profile of other funds' signals, which greatly complicates the analysis. Further, fund expenses are empirically much more persistent than performance, and recent evidence shows that the fees charged for the management of a given fund are only rarely changed (Kuhnen 2004). This evidence suggests that not much is lost, and much is won in terms of tractability by assuming that fees are set prior to observing the performance signal.

¹⁰ Given the empirical literature on performance persistence, $\nu = 0.55$ is a plausible value for the precision of the signal and $\nu = 0.7$ can be considered as an upper bound for this precision. The value $\nu = 0.9$ is analyzed as a limiting case.

5 Evidence on the Relationship between Performance and Fees

5.1 Mutual Fund Fee Structure

In the market for mutual funds, the fees paid by investors take two forms: periodic fees (operating expenses) and one-time fees (loads).¹¹ Expenses mostly consist of management fees, but also include 12b-1 (distribution and marketing) fees, custody fees, and administrative fees, as well as operating, legal, and accounting costs. They are computed as a percentage of assets under management (termed the *expense ratio*) and are deducted on a daily basis from the fund's net assets by the managing company. Fees paid to brokers in the course of the fund's trading activity are not included in the fund's expense ratio.

Loads are generally used to pay distributors, and they differ from operating expenses in that they are paid by the individual investor as a fraction of the amount invested at the time of purchasing fund shares (*sales charge on purchases*) or redeeming fund shares (*deferred sales charge*). Since fund returns are typically computed from the fund's net asset value, quoted returns are net-of-expenses, but before loads.

5.2 Empirical Evidence

The quality of an actively managed fund is commonly defined as the manager's ability to deliver returns above those that any investor could obtain following a passive strategy, such as investing in an index fund. Differences in managerial quality could translate into differences in after-expense returns if quality were not fully priced. If higher quality were partly priced (that is, if better funds charged higher expenses) differences in after-expense returns would be smaller than differences in fees. On the other hand, if high-quality funds happened to charge lower fees, differences in after-expense returns would be greater than differences in fees.

Elton et al. divide a sample of U.S. mutual funds available in the 1965-84 period into quintiles by expense ratios and measure average after-expense risk-adjusted returns¹² for funds in each quintile. They find that funds with higher

¹¹ Mahoney (2004) provides a review of mutual fund fee practices and regulation. For a more detailed description, we suggest that the reader visit the Online Publications section of the SEC internet site.

¹² In particular, they measure performance as the intercept term from the regression of annual returns on three indexes, tracking the evolution of stocks in the S&P 500 index, non-S&P stocks, and bonds. This measure of performance can be interpreted

expense ratios perform significantly worse, and that performance differences between funds in the best and the worst quintiles exceed differences in fees, which suggests that funds with higher expenses exhibit lower before-expense returns. Put differently, low-quality funds seem to be more expensive. Gruber studies cross-sectional differences in after-fee performance in the 1985-94 period and finds similar results.¹³

Harless and Peterson analyze data employed in early performance studies and find that the predictive power of expense ratios with respect to performance also extends to the 1954-64 period. They conclude that “all the studies show that funds with the lowest expense ratios tend to perform best, and funds with highest expense ratios tend to perform worst” (p. 262).

Carhart proposes a different measure of performance.¹⁴ When regressing this performance measure on expense ratios, he estimates that in the 1962-93 period funds with annual expenses of 100 basis points above the average had on average 154 basis points below mean after-expense performance. Again, the effect of fees is to amplify rather than mitigate differences in before-expense performance.

Finally, Chevalier and Ellison, using a measure of performance similar to Carhart’s, report that manager and fund characteristics (such as the portfolio turnover ratio and log of assets) contribute to explaining differences in performance in the 1988-95 period. When controlling for these variables, they provide estimates of the effect on after-expense performance of a 100 basis point reduction in expense ratios that range from 152 to 225 basis points. Otten and Bams (2002) and Bechmann and Rangvid (2007) have recently reported evidence of a negative relation between expenses and fund performance in several European countries.

Put together, the empirical evidence implies that superior management is not priced through higher expense ratios. On the contrary, it appears that the effect of expenses on after-expense performance (even after controlling for funds’ observable characteristics) is more than one-to-one, which would imply that low-quality funds charge higher fees. Price and quality thus seem to be inversely related in the market for actively managed mutual funds.

as the value added by the fund manager with respect to a passive strategy.

¹³ Gruber analyzes three alternative proxies for performance: (i) the fund’s average return relative to the market, (ii) the fund’s average return in excess of the fund’s expected return according to the Capital Asset Pricing Model, and (iii) the fund’s average excess return according to a four-index model.

¹⁴ Carhart employs a four-factor model that captures the fund’s exposure to sources of undiversifiable risk.

6 Conclusion

In this paper, we have shown that in the mutual fund industry, better-quality funds should not be expected to charge higher prices. Moreover, investors' limited ability to evaluate fund quality may lead to equilibria in which worse-performing funds charge higher fees. We thus obtain a form of inverse price differentiation that is consistent with existing evidence on mutual fund performance.

The fundamental role played by mutual funds and the current demands for regulatory action call for further analysis of the mutual fund industry. Our model suggests several directions for future research. In particular, a richer model of mutual fund competition should capture a wider set of strategic choices by mutual fund managers. For instance, while we have taken fund quality as exogenous, mutual fund management companies may, to some extent, set the quality of the funds they offer through their choice of managers or their expenditure in market analysis.

Also, funds might try to differentiate themselves through their marketing decisions. One could expect that old-enough funds with a good performance record could advertise that performance record to attract unsophisticated investors. Moreover, the theoretical literature in industrial organization offers additional rationales for the existence of informative advertising (for a recent review of the literature, see Bagwell 2007). In the context of the mutual fund sector, however, our conjecture is that advertising could well be associated with lower rather than with higher quality. In particular, it may not be difficult for low-quality funds to advertise on the basis of some measure of performance (such as short term risk-unadjusted returns). Moreover, it is not clear that it is in the interest of better funds to advertise. The reason is that while unsophisticated investors are likely to respond to advertising, sophisticated investors can be expected to be insensitive to advertising, since they can obtain and process publicly available information about fund performance. Therefore, funds catering only to the unsophisticated segment might want to advertise in order to gain market share in that sector, while better funds might avoid advertising precisely to distinguish themselves from bad funds. Consistently with this reasoning, Gallaher et al. (2006, table 8) have recently found that mutual funds with better performance do not have larger advertising expenditures, and, for some types of funds, larger advertising expenditures are associated with lower performance.

Although in this paper we have specifically modelled the market for mutual funds, the insights of the model may be generalizable to other markets where quality assessment is costly. The existing empirical evidence on the relationship between quality and price suggests that the correlation between these two

variables is typically not strongly positive and that in a significant number of markets, the correlation is indeed negative (e.g., Caves and Greene 1996). In our model, the negative correlation between price and quality results from the combination of asymmetric information about product quality and the presence of a subset of unsophisticated investors. Future work may further explore, both theoretically and empirically, how these and other factors affect observed price-quality correlations.

Our results indicate that the complexity associated with the evaluation of fund quality may, on the one hand, weaken competition, leading to high average fees even in the presence of a large pool of competing mutual funds, and, on the other hand, lead to a segmented market in which a fraction of investors pay higher than average fees for underperforming funds. Whether this state of things could be improved by regulation and what the optimal form of this regulation is are questions that merit further scrutiny. Currently, both the SEC and NASD, the self-regulatory arm of the U.S. securities industry, impose limits on redemption fees and the loads that can be charged to mutual fund investors to pay for brokerage services. Further research is, however, required to assess the potential effects of similar limits to mutual fund expenses. For example, if there were economies of scale in fund management, regulatory pressure to reduce fees could disproportionately affect young, small funds. This could have an adverse impact on competition in the sector by restricting entry. A less contentious alternative, already having been pursued by the SEC, is to require that funds improve the disclosure of their fees to allow investors to realize the dollar cost of the expenses paid. The SEC, however, has stopped short of requiring funds to disclose periodically to each investor the exact dollar amount of the expenses paid on the grounds that such a requirement would impose large processing costs on mutual fund companies. Given the large potential costs that a poor understanding of the impact of fees on returns has for unsophisticated investors, our model would suggest reevaluating this cost-benefit analysis. Our model also suggests that requiring funds to disclose the level of fees charged by the fund compared to the average or median fees in the corresponding investment category, information that is already voluntarily disclosed by some mutual fund management companies, could greatly contribute to prevent funds from overcharging unsophisticated investors.

Acknowledgements

The authors thank the Editor, the Associate Editor, two anonymous referees, Sandro Brusco, Luis Corchón, María Gutiérrez, and Gonzalo Rubio, as well as seminar participants at Universitat Pompeu Fabra, Universitat de les Illes Balears, Econometric Society European Meetings (2004), and EARIE

(2005) for their comments and suggestions. The usual disclaimer applies. The financial support of Spain's Ministry of Education and Science (SEJ 2004-01668/ECON and SEJ2005-06655/ECON) and the BBVA Foundation (Ruiz-Verdú) is gratefully acknowledged.

References

- [1] Alexander, G.J., Jones, J.D., Nigro, P.J., 1997. Mutual fund shareholders: Characteristics, investor knowledge, and sources of information. Office of the Comptroller of the Currency, E&PA Working Paper 97-13.
- [2] Bagwell, K., 2007. The economic analysis of advertising. In: Armstrong, M., Porter, R. (Eds.). *Handbook of Industrial Organization*, Vol. 3. Amsterdam: North-Holland, 1701-1844.
- [3] Bagwell, K., Riordan, M.H., 1991. High and declining prices signal product quality. *American Economic Review* 81, 224-239.
- [4] Barber, B.M., Terrance, O., Zheng, L., 2005. Out of sight, out of mind: the effects of expenses on mutual fund flows. *Journal of Business* 78, 2095-2120.
- [5] Bechmann, K.L., Rangvid, J., 2007. Rating mutual funds: Construction and information content of an investor-cost based rating of Danish mutual funds. *Journal of Empirical Finance* 14, 662-693.
- [6] Berk, J.B., Green, R.C., 2004. Mutual fund flows and performance in rational markets. *Journal of Political Economy* 112, 1269-1295.
- [7] Bhattacharya, S., Pfleiderer, P., 1985. Delegated portfolio management. *Journal of Economic Theory* 36, 1-25.
- [8] Bollen, N.P.B., Busse, J.A., 2005. Short-term persistence in mutual fund performance. *Review of Financial Studies* 18, 569-597.
- [9] Capon, N., Fitzimmons, G., Prince, R., 1996. An individual level analysis of the mutual fund investment decision. *Journal of Financial Services Research* 10, 59-82.
- [10] Carhart, M., 1997. On persistence in mutual fund performance. *Journal of Finance* 52, 57-82.
- [11] Caves, R.E., Greene, D.P., 1996. Brands' quality levels, prices, and advertising outlays: empirical evidence on signals and information costs. *International Journal of Industrial Organization* 14, 29-52.
- [12] Chevalier, J., Ellison, G., 1999. Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance. *Journal of Finance* 54, 875-899.
- [13] Christoffersen, S., Musto, D., 2002. Demand curves and the pricing of money management. *Review of Financial Studies* 15, 1499-1524.
- [14] Cox, J.D., Payne, J., 2005. Mutual fund expense disclosures: a behavioral perspective. *Washington University Law Quarterly* 84, 907-938.
- [15] Das, S.R., Sundaram, R.K., 2002. Fee speech: Signaling, risk-sharing, and the impact of fee structures on investor welfare. *Review of Financial Studies* 15, 1465-1497.
- [16] Dwyer, P., Borras, A., Young, L., 2003. Breach of trust. *Business Week*. New York: Dec 15. p. 98.
- [17] *The Economist*, 2004. Low fidelity. London: Sep 25. Vol. 372, Iss. 8394, p. 106.
- [18] Elton, E.J., Gruber, M.J., Das, S., Matthew, H., 1993. Efficiency with

- costly information: a reinterpretation of evidence from managed portfolios. *Review of Financial Studies* 6, 1–22.
- [19] Fluet, C., Garella, P.G., 2002. Advertising and prices as signals of quality in a regime of price rivalry. *International Journal of Industrial Organization* 20, 907–930.
 - [20] Gabaix, X., Laibson, D., 2003. Some industrial organization with boundedly rational consumers. Working paper, Harvard University.
 - [21] Gabaix, X., Laibson, D., 2006. Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics* 121, 505–540.
 - [22] Gallaher, S., Kaniel, R., Starks, L., 2006. Madison Avenue meets Wall Street: mutual fund families, competition and advertising. Working paper, Duke University-Fuqua School of Business.
 - [23] General Accounting Office, 2000. Mutual fund fees. Additional disclosure could encourage price competition (GAO/GGD-00-126). Washington D.C.: General Accounting Office.
 - [24] Grinblatt, M., Titman, S., 1992. The persistence of mutual fund performance. *Journal of Finance* 47, 1977–1984.
 - [25] Gruber, M., 1996. Another puzzle: The growth in actively managed mutual funds. *Journal of Finance* 52, 783–810.
 - [26] Harless, D.W., Peterson, S.P., 1998. Investor behavior and the persistence of poorly-performing mutual funds. *Journal of Economic Behavior and Organization* 37, 257–276.
 - [27] Hendricks, D., Patel, J., Zeckhauser, R., 1993. Hot hands in mutual funds: short run persistence of relative performance, 1974–1988. *Journal of Finance* 48, 93–130.
 - [28] Hertzendorf, M.N., Overgaard, P.B., 2001. Price competition and advertising signals: Signaling by competing senders. *Journal of Economics and Management Strategy* 10, 621–662.
 - [29] Hortacsu, A., Syverson, C., 2004. Product differentiation, search costs, and the welfare effects of entry: a case study of S&P 500 index funds. *Quarterly Journal of Economics* 119, 403–456.
 - [30] Investment Company Institute, 2006. Mutual Fund Factbook, 46th Edition. Washington, D.C.: Investment Company Institute.
 - [31] Kuhnen, C.M., 2004. Dynamic contracting in the mutual fund industry. Working paper, Stanford Graduate School of Business.
 - [32] Mahoney, P.G., 2004. Manager-investor conflicts in mutual funds. *Journal of Economic Perspectives* 18, 161–182.
 - [33] Metrick, A., Zeckhauser, R., 1999. Price versus quantity: Market clearing mechanisms when sellers differ in quality. *Journal of Risk and Uncertainty* 17, 175–190.
 - [34] Nanda, V., Narayanan, M.P., Warther, V.A., 2000. Liquidity, investment ability, and mutual fund structure. *Journal of Financial Economics*, 57, 417–443.
 - [35] Otten, R., Bams, D., 2002. European mutual fund performance. *European*

- Financial Management 8, 75–101.
- [36] Palomino, F., Prat, A., 2003. Risk taking and optimal contracts for money managers. *RAND Journal of Economics* 34, 113–137.
 - [37] Salop, S., Stiglitz, J.E., 1977. Bargains and ripoffs: A model of monopolistically competitive price dispersion. *Review of Economic Studies* 44, 493–510.
 - [38] Securities and Exchange Commission, 2000. Report on mutual fund fees and expenses. Washington, D.C.: Securities and Exchange Commission, Division of Investment Management.
 - [39] Securities and Exchange Commission, 2004. Disclosure regarding portfolio managers of registered management investment companies, 69 Federal Register, August 27.
 - [40] Schmalensee, R., 1978. A model of advertising and market structure. *Journal of Political Economy* 86, 485–503.
 - [41] Shaked, A., Sutton, J., 1982. Relaxing price competition through product differentiation. *Review of Economic Studies* 49, 3–13.
 - [42] Slovic, P., 1972. From Shakespeare to Simon: Speculations and some evidence about man's ability to process information. Oregon Research Institute Monograph, 12.
 - [43] Wermers, R., 2000. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *Journal of Finance* 55, 1655–1695.
 - [44] Wilcox, R.T., 2003. Bargain hunting or star gazing? Investors' preferences for stock mutual funds. *Journal of Business* 76, 645–663.

Figures

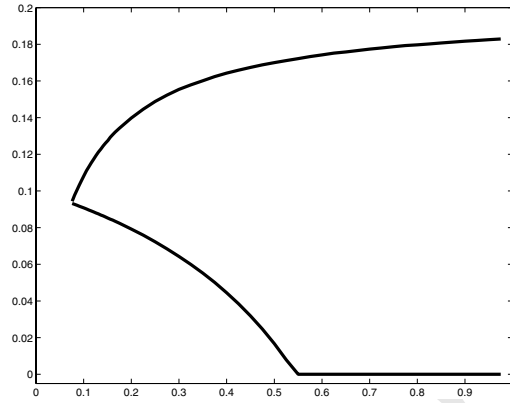


Fig. A.1. Conditions for Existence of Separating Equilibria with Unsophisticated Investors. $R_b = 1.1$; $R_g = 1.3$; $p = 0.5$; $N=2$. The x-axis displays values of γ , while c is displayed along the y-axis. The curves represent the minimum and maximum values of c such that a separating equilibrium exists for each γ .

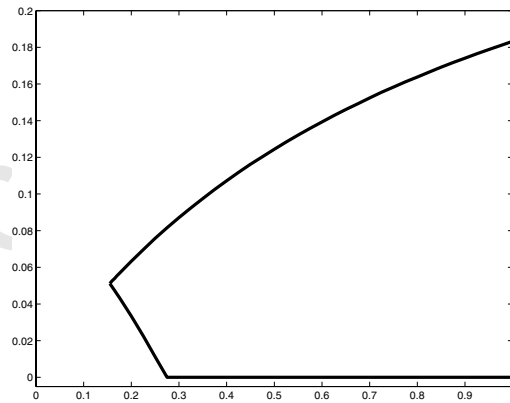


Fig. A.2. Conditions for Existence of Separating Equilibria with Unsophisticated Investors. $R_b = 1.1$; $R_g = 1.3$; $p = 0.5$; $N=5$. The x-axis displays values of γ , while c is displayed along the y-axis. The curves represent the minimum and maximum values of c such that a separating equilibrium exists for each γ .

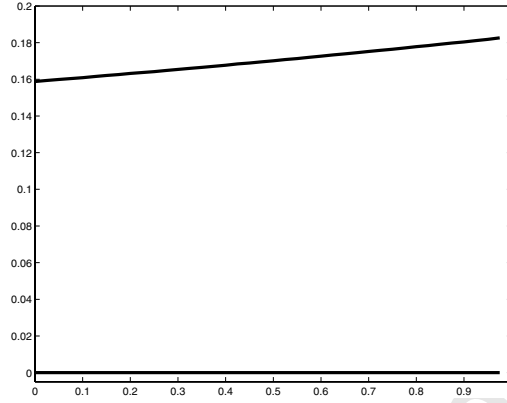


Fig. A.3. Conditions for Existence of Pooling Equilibria with Unsophisticated Investors. $R_b = 1.1$; $R_g = 1.3$; $p = 0.5$; $N=2$. The x-axis displays values of γ , while c is displayed along the y-axis. The curves represent the minimum and maximum values of c such that a pooling equilibrium exists for each γ .

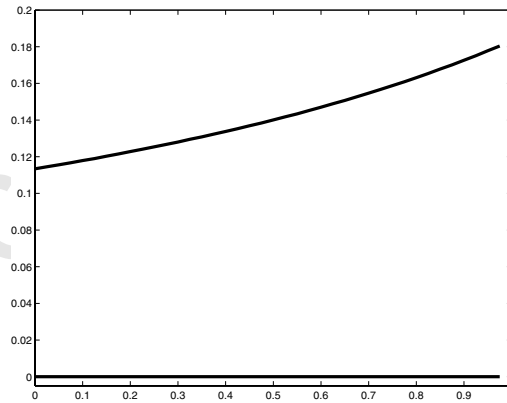


Fig. A.4. Conditions for Existence of Pooling Equilibria with Unsophisticated Investors. $R_b = 1.1$; $R_g = 1.3$; $p = 0.5$; $N=5$. The x-axis displays values of γ , while c is displayed along the y-axis. The curves represent the minimum and maximum values of c such that a pooling equilibrium exists for each γ .

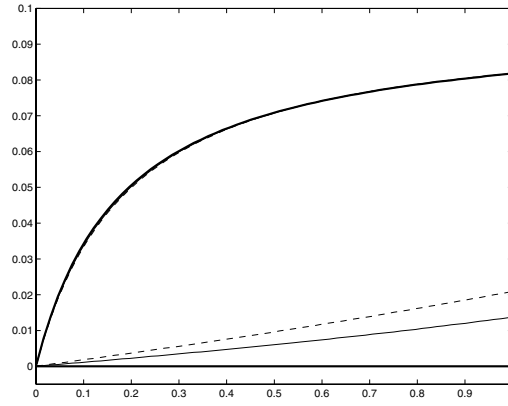


Fig. A.5. Conditions for Existence of Separating Equilibria with Unsophisticated Investors. $R_b = 0.9$; $R_g = 1.3$; $p = 0.5$; $\nu = 0.55$; $N = 2$. The x-axis displays values of γ , while c is displayed along the y-axis. The curves represent the minimum and maximum values of c such that a separating equilibrium exists for each γ . We have considered three possible reactions of unsophisticated investors to observed differences in quality signals: (i) $f_h(\sigma_l) = f_l(\sigma_h) = \frac{1}{2}$; (ii) $f_h(\sigma_l) = 0.8, f_l(\sigma_h) = 0.2$; and (iii) $f_h(\sigma_l) = 1, f_l(\sigma_h) = 0$. In all cases $f_h(\sigma_h) = f_l(\sigma_l) = \frac{1}{2}$. Situations (i)-(iii) are represented by thick solid lines, thin solid lines and dashed lines, respectively. Note that the lines may overlap, so that only the thick solid line is visible.

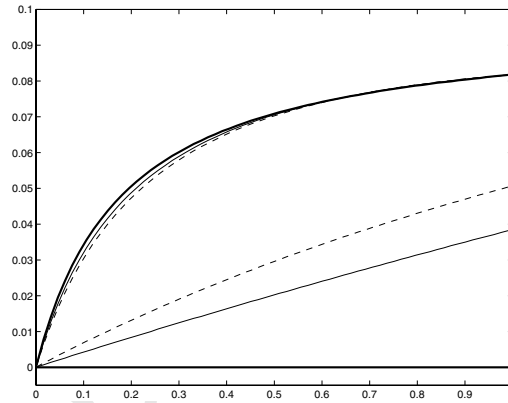


Fig. A.6. Conditions for Existence of Separating Equilibria with Unsophisticated Investors. $R_b = 0.9$; $R_g = 1.3$; $p = 0.5$; $\nu = 0.7$; $N = 2$. The x-axis displays values of γ , while c is displayed along the y-axis. The curves represent the minimum and maximum values of c such that a separating equilibrium exists for each γ . We have considered three possible reactions of unsophisticated investors to observed differences in quality signals: (i) $f_h(\sigma_l) = f_l(\sigma_h) = \frac{1}{2}$; (ii) $f_h(\sigma_l) = 0.8, f_l(\sigma_h) = 0.2$; and (iii) $f_h(\sigma_l) = 1, f_l(\sigma_h) = 0$. In all cases $f_h(\sigma_h) = f_l(\sigma_l) = \frac{1}{2}$. Situations (i)-(iii) are represented by thick solid lines, thin solid lines and dashed lines, respectively.

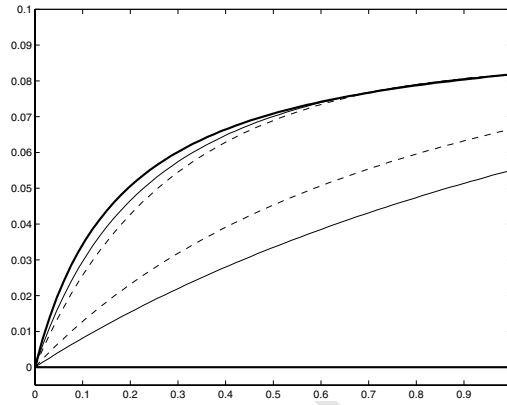


Fig. A.7. Conditions for Existence of Separating Equilibria with Unsophisticated Investors. $R_b = 0.9$; $R_g = 1.3$; $p = 0.5$; $\nu = 0.9$; $N = 2$. The x-axis displays values of γ , while c is displayed along the y-axis. The curves represent the minimum and maximum values of c such that a separating equilibrium exists for each γ . We have considered three possible reactions of unsophisticated investors to observed differences in quality signals: (i) $f_h(\sigma_l) = f_l(\sigma_h) = \frac{1}{2}$; (ii) $f_h(\sigma_l) = 0.8, f_l(\sigma_h) = 0.2$; and (iii) $f_h(\sigma_l) = 1, f_l(\sigma_h) = 0$. In all cases $f_h(\sigma_h) = f_l(\sigma_l) = \frac{1}{2}$. Situations (i)-(iii) are represented by thick solid lines, thin solid lines and dashed lines, respectively.